

Archaeogenomic analysis of the first steps of Neolithisation in Anatolia and the Aegean

Gülşah Merve Kılınç¹, Dilek Koptekin^{2*}, Çiğdem Atakuman^{3*}, Arev Pelin Sümer⁴, Handan Melike Dönertaş⁵, Reyhan Yaka⁴, Cemal Can Bilgin⁴, Ali Metin Büyükkarakaya⁶, Douglas Baird⁷, Ezgi Altınışik⁸, Pavel Flegontov^{8,9,10}, Anders Götherström^{1,#}, İnci Togan^{4,#}, Mehmet Somel^{4,#}

Affiliations:

¹Department of Archaeology and Classical Studies, Stockholm University, Lilla Frescativaegen 7, 114 18, Stockholm, Sweden

²Department of Health Informatics, Middle East Technical University, 06800, Ankara, Turkey

³Department of Settlement Archaeology, Middle East Technical University, 06800, Ankara, Turkey

⁴Department of Biological Sciences, Middle East Technical University, 06800, Ankara, Turkey

⁵European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SD, United Kingdom.

⁶Department of Anthropology, Hacettepe University, Beytepe, 06800, Ankara, Turkey

⁷Department of Archaeology, Classics, and Egyptology, Liverpool University, Liverpool L69 WZ, United Kingdom

⁸Department of Biology and Ecology, Faculty of Science, University of Ostrava, Ostrava, Czech Republic

⁹A.A. Kharkevich Institute for Information Transmission Problems, Russian Academy of Sciences, Moscow, Russia

¹⁰Institute of Parasitology, Biology Centre, Czech Academy of Sciences, České Budějovice, Czech

*equal contribution

#co-senior authors

Abstract

The Neolithic transition in west Eurasia occurred in two main steps: the gradual development of sedentism and plant cultivation in the Near East, and the subsequent spread of Neolithic cultures into the Aegean and across Europe after 7,000 cal BCE. Here we use published ancient genomes to investigate gene flow events in west Eurasia during the Neolithic transition. We confirm that the early Neolithic central Anatolians in the ninth millennium BCE were likely descendants of local hunter-gatherers, rather than immigrants from the Levant or Iran. We further study the emergence of post-7,000 cal BCE north Aegean Neolithic communities. Although Aegean farmers have frequently been assumed to be colonists originating from either central Anatolia or from the Levant, our findings raise alternative possibilities: north Aegean Neolithic populations may have been the product of multiple westward migrations, including south Anatolian emigrants, or they may have been descendants of local Aegean Mesolithic groups who adopted farming. These scenarios are consistent with the diversity of material cultures among Aegean Neolithic communities and the inheritance of local forager know-how. The demographic and cultural dynamics behind the earliest spread of Neolithic culture in the Aegean could therefore be distinct from the subsequent Neolithization of mainland Europe.

Keywords

Ancient DNA, archaeogenomics, Neolithic, migration, acculturation, population genetics

1. Introduction

The primary zone of Neolithisation in western Eurasia encompassed the Levant, Taurus-Zagros ranges of Mesopotamia, central Anatolia and Cyprus [1–4]. The earliest evidence for sedentary life and food storage in this region goes back to the Natufians (c.12,500–10,800 cal BCE) [5,6]. Sedentary communities were established across this zone during the first phase of the Pre-Pottery Neolithic (PPN, or Aceramic Neolithic, c.10,000-8,500 cal BCE), and the first indications of plant cultivation appeared [7–9]. Between c.8,500–7,000 cal BCE, community sizes increased, architectural elaboration intensified, and a subsistence economy based on agriculture gradually became the norm [10–14]. Meanwhile, portable artifacts such as figurines and stamps evolved into staples of sedentary life, and pottery production became widespread around c.7,000 cal BCE [10,11]. The elements of the subsequent Pottery Neolithic culture (PN, c.7,000-5,500 cal BCE), including integrated cultivation practices of domestic plants and animals, the architectural practices of sedentary life, together with portable artifacts, have been collectively described as the Near Eastern “Neolithic Package” [15–18].

During the same period, there were no signs of a Neolithisation in west Anatolia and the Aegean. Only after c.7,000 cal BCE did elements of the “Neolithic Package” appear in these regions, eventually spreading toward Europe [19–21]. Some archaeologists suggest that the emergence of the Neolithic elements in the Aegean and in Europe without a preceding PPN development period indicates the role of demic processes, i.e. migrations from the Neolithic primary zones through land and sea routes, frequently described as a leap-frog model where migrants form enclaves in new territory [15, 16, 22-27]. Others, in contrast, favour a role for interaction between local foragers and primary zone Neolithic populations, including the adoption of Neolithic elements by locals and acculturation [16, 28-30].

Recent archaeogenomic data has shown that the Neolithization of central, western and northern Europe involved migration from a single eastern source, frequently termed “Anatolian farmer” [31-34], while in other regions, such as in the Baltic [34,35] and in South Greece [34], acculturation may have played role.

In most of Europe, there is limited genetic evidence for early admixture between farmers and local European Mesolithic (WHG) communities in the 6th millennium BCE, such that early European farmers studied to date (with few exceptions [36]) carry ancestry similar to farmers from northwest Anatolian Barcin [32-34] (**electronic supplementary material, figure S1**). In subsequent millenia, however, WHG-like ancestry appears in middle and late Neolithic European populations [37-39]. These observations support a leap-frog model of Neolithic spread in Europe [28]: farmers only occupied enclaves in the new territories while Mesolithic groups persisted in the same regions [40-43].

The processes behind the earliest steps of Neolithization and the Neolithic spread in the Aegean are less understood. For instance, whether Aegean Neolithic populations were recent colonists originating from areas of the primary zone of Neolithisation (e.g. [27]), descendants of indigenous foragers (e.g. [20]), or admixed groups (e.g. [15,16]) is still contentious. Additionally, whether Aegeans' demographic or cultural relationships were stronger with central Anatolians [16] or with Levantine seafaring populations [27] remains unclear. We re-analyse published ancient human genomes to answer these questions and to dissect the demographic dynamics behind the Neolithic transition in Anatolia and the Aegean.

2. Methods

(a) Compiling and mapping genomic data

We obtained DNA sequencing data of 99 published ancient individuals (**electronic supplementary material, table S1**), generated using either whole genome shotgun sequencing and/or sequencing of libraries enriched by hybridization capture [31-33, 36, 38, 39, 43-47]. We mapped sequencing reads to the human reference genome (hs37d5) using the Burrows-Wheeler Aligner (BWA, v.0.7.12) [48], with the parameters “-l 16500, -n 0.01, -o 2”. We filtered PCR duplicates using FilterUniqSAMCons.py [49]. We

95 filtered reads shorter than 35 base pairs, with >10% mismatches to the reference, and <30 mapping quality
96 per read.

98 **(b) Preparation of population genetics analysis data sets**

99
100 We restricted our analysis to known present-day DNA variants to minimize false positives. We used two
101 different modern reference panels, calling genotypes of ancient individuals for SNPs overlapping with (i)
102 the Human Origins genotype dataset [43,50] and (ii) the 1000 Genomes whole genome sequence data [51]
103 using SAMtools *mpileup* (v.1.3) [52]. For (i) we obtained a curated version of the Human Origins panel of
104 594,924 autosomal SNP genotype calls for 2,730 present-day individuals from [43]. We determined the
105 SNPs of the ancient samples overlapping with this dataset. We encoded transitions as missing to avoid
106 confounding with cytosine deamination in ancient DNA. To prepare (ii) we obtained the BAM and VCF
107 files for the African Yoruba individuals from 1000 Genomes Project phase 3 from [51]. Using vcftools
108 [53], we extracted a total of 1,938,919 transversion SNPs with minor allele frequencies of $\geq 10\%$ in the
109 Yoruba population to avoid false positive calls [37, 41]. We determined the positions in the ancient samples
110 overlapping with this dataset. We merged ancient genotypes with these two datasets using PLINK [54]
111 requiring base quality ≥ 30 per overlapping position. We haploidised each full dataset by randomly selecting
112 one allele per position. The Human Origins-merged dataset, which has higher number of present-day
113 populations, was used for principal component analysis and for calculating f_3 -statistics. The 1000 Genomes-
114 merged dataset, with a higher number of SNPs, was used for D -statistics, where we require high statistical
115 power.

117 **(c) Principal component analysis**

118
119 We performed principal component analysis (PCA) by calculating principal components using west
120 Eurasian populations from the Human Origins dataset using the *smartpca* program of EIGENSOFT [55]

with the “numoutlieriter:0” parameter. We projected ancient genomes onto the reference space using the “lsqproject:YES” option and plotted the results using R (v.3.3.0).

(d) D - and f_3 -statistics

We computed D -statistics using the *qpDstat* program of ADMIXTOOLS package [50]. We assessed statistical significance by calculating standard errors using a block jackknife of 0.5 Mbp. We used the Yoruba population as outgroup for the D -statistics [33]. We computed f_3 -statistics, *i.e.* genetic affinity between pairs of populations based on an estimate of shared drift between them since their divergence from an outgroup population, using the *qp3Pop* program of the ADMIXTOOLS package [50]. The Human Origins data set’s African Mbuti population was used as outgroup for calculating f_3 -statistics [43]. We performed multiple testing correction using Benjamini-Yekutieli method for all 207 D -statistics results and reported adjusted p -values together with Z scores per each test [56]. For the pairwise f_3 -statistics, as genetic distance measure between a pair of populations, X and Y, we used: $1 - f_3(\text{Mbuti}; X, Y)$ [31]. These pairwise distances were summarized with the multidimensional scaling (MDS) method using the *cmdscale* function of R. We evaluated the goodness of fit for MDS using “GOF” component obtained from *cmdscale* function.

(e) Heterozygosity estimates

We calculated heterozygosity as a measure of genetic diversity in a population, using genome sequence data of (i) Bon002 (from Boncuklu, central Anatolia, pre-7,000 cal BCE) [31], Tep003 (Tepecik-Çiftlik, central Anatolia, post-7,000 cal BCE) [31], Bar8 (Barcın, north Aegean, post-7,000 cal BCE) [32], and Rev5 (Revenia, north Aegean, post-7,000 cal BCE) [32]. We calculated genome coverage per sample using GenomeCoverageBed [57]. We downsampled the genome sequences of Bon002 and Bar8 to similar levels as the other two samples using SAMtools (v.1.3) [52]. We calculated heterozygosity per sample using ANGSD [58] as “angsd -GL 1 -doGlf 2 -doMajorMinor 1 -sites ReferenceSNP.pos -bam bamlist -doSaf 1

-anc referencegenome.fasta”. To minimise false positives, we only considered transversions overlapping with of Yoruba individuals from 1000 Genomes Project phase 3 from [51].

(f) Modelling of admixture

We used the *qpWave/qpAdm* framework [39, 59] in the AdmixTools package [50] to model populations as mixtures of two or more sources. The following worldwide set of ancient and present-day outgroups, which most probably did not experience any post-split gene flow from Anatolian/Aegean populations, was used: Mbuti, Yoruba, Ust Ishim, El Miron, Goyet Q116, Villabruna, Kostenki14, Vestonice16, Papuan, Onge, Karitiana, Mixe, Chipewyan, Oroqen, Koryak, Dai, Japanese. Adding East European hunter-gatherers (EHG) as a close outgroup to increase the resolution did not change the results.

(g) Serial coalescent simulations

We performed serial coalescent simulations using fastsimcoal [60] under 4 various demographic models involving Neolithic central Anatolians, Aegeans, Iranians, and WHG (not including Levantine populations, for whom we lack whole genome data). The simulations were designed to mimick the data with respect to tree topology, divergence times, and sample sizes. We then performed *D*-statistics on the simulated DNA and compared these with the observed data to gain understanding into the plausibility of different models. Specifically, we generated data to represent Iranian Neolithics (10,000 BP), WHGs (Loschbour: 7,200 BP), central Anatolian Neolithics (Tepecik-Çiftlik: 8,500 BP; Boncuklu: 10,000 BP), the Aegean Neolithics (Revenia: 8,300 BP) and present-day sub-Saharan Africans (Yoruba-YRI). We launched 100 runs for each model defined in the paramater file (input.par) for testing different population histories. For all models, we sampled 30 Mb DNA sequences for: 5 present-day Yoruba, 2 Iranian Neolithics, 2 WHGs, 4 central Anatolian Neolithics (2 Tepecik-Çiftlik, 2 Boncuklu) and 2 Aegean Neolithics (Revenia). We assumed a mutation rate of 1.00E-09 bp/year, and a recombination rate of 1.00E-08 bp/year, and assumed 25 years per

generation, again following [46]. We set the effective population size (N_e) of these populations and times of divergence between Anatolian Neolithic, WHGs and Iranian Neolithic populations based on [46]. We converted all outputs (arp file) to plink format and computed D -statistics with topology of $D(YRI, Test, central\ Anatolian\ N, Aegean\ N)$ to test the relationships among populations via AdmixTools [50]. Note that the tree topology involving the Anatolian/Aegean populations, Iran, WHG, and the Africans, were based on the phylogenetic analysis from [46]. The Anatolian/Aegean populations were assumed to diverge simultaneously from the same source (star-shaped).

3. Results

(a) Early Holocene gene pools of west Eurasia and the Anatolian/Aegean gene pool

We compiled published genome sequence data of 99 ancient individuals (sample ages: c.11,840-4,360 cal BCE) (**electronic supplementary material, figure S2, table S1**). Both a PCA using present-day and ancient populations (**electronic supplementary material, figure S1**) and an MDS analysis using only ancient genomes (**figure 1,2, electronic supplementary material, table S2**) revealed the presence of four distinct gene pools in early Holocene west Eurasia: (a) a “Caucasia/Iran gene pool”, (b) a “Levant gene pool”, (c) a “European pre-Neolithic gene pool”, (d) an “Anatolian/Aegean gene pool”. To objectively measure clustering in gene pools a-c, we used D -statistics of the form $D(Yoruba, p1; p2, p3)$ where “ p ” refers to the Caucasia/Iran, the Levant, or European pre-Neolithic gene pools, correcting for multiple testing. In 80% comparisons ($p < 0.05$; $Z \geq 3$), populations belonging to the same gene pool shared more alleles with each other compared to external populations (**figure 1a, electronic supplementary material, figure S3a, table S3**). The only exceptions were comparisons involving a single pre-Neolithic individual from Iran for which we had relatively few SNPs and low statistical power.

We then investigated the relationships among ancient Anatolians and other west Eurasian gene pools, using the oldest Anatolian population yet sequenced: Boncuklu from central Anatolia (sample ages: c.8,300-7,952 cal BCE), an Aceramic Neolithic population previously predicted to be the descendants of local Epipaleolithic groups [31, 61]. We computed D -statistics of the form $D(\text{Yoruba}, p1; p2, \text{Boncuklu})$, where “ $p1$ ” and “ $p2$ ” refer to populations belonging to different gene pools: Caucasias/Iran, the Levant, or the European pre-Neolithic. In 56% of the comparisons ($p < 0.05$; $Z \geq 2.8$), all three regional gene pools showed higher affinity to Boncuklu than to each other (figure 1a-1b, electronic supplementary material, table S4). Using the qpWave/qpAdm algorithm [39, 59] we further modelled the Boncuklu population as a mixture of CHG (59.1%), Levant (31.4%) and WHG (9.5%) (electronic supplementary material, table S5).

We next included three post-7,000 cal BCE Neolithic populations from Anatolia and Aegean in the analyses: Tepecik-Çiftlik in central Anatolia [31,62], Barcın in northwest Anatolia [32, 38, 63], and “Revenia” in Pieria of northeast Greece [32]. We computed D -statistics of the form $D(\text{Yoruba}, p1; p2, p3)$ where “ $p1$ ” and “ $p2$ ” are Anatolian/Aegean populations and “ $p3$ ” is an external population (Caucasias/Iran, Levant, or European pre-Neolithic). In 94% of the comparisons ($p < 0.05$; $Z \geq 2.8$) all Anatolian/Aegean populations were genetically closer to each other than to any other gene pool (electronic supplementary material, figure S3b, table S5).

Given archaeological indication that Aegean Neolithic was influenced by east Mediterranean sources [27], we further studied the genetic affinities of Aegean Neolithic people to central Anatolian Neolithics and to the Levantines. Calculating D -statistics of the form $D(\text{Yoruba}, \text{northAegean}; \text{Levant}, \text{centralAnatolia})$ revealed that the post-7,000 cal BCE Neolithic north Aegean individuals (Barcın and Revenia) consistently share more alleles with central Anatolians compared to south Levantines, where 50% of the comparisons were significant ($p < 0.05$; $Z \geq 2.8$) (electronic supplementary material, figure S3b, table S5).

(b) Notable genetic diversity in the Aegean

To assess demographic events in the Near East during the Neolithic transition we studied signatures of regional admixture using diachronic populations from the same region (**figure 1c**). In 83% of the comparisons, pre-7,000 cal BCE Neolithic populations of the Levant and of Iran were genetically closer to all post-7,000 cal BCE Anatolian/Aegean populations (Tepecik, Barcin, Revenia) compared to the 7,000 cal BCE Anatolian Boncuklu ($p < 0.05$; $Z \geq 3$) (**electronic supplementary material, figure S3c, table S6**). Considering the radiocarbon dates of the investigated individuals, this is consistent with gene flow from both the Levant and from Iran into Anatolia, within a period ranging from the PPN to the PN (**figure 1c, arrows “e” and “f”**). These results are also compatible with a regional increase in the levels of admixture during the Neolithic [31, 33], although alternative explanations to gene flow remain plausible, such as population structure confounding the analysis results [64].

Next, to gain understanding into Aegean Neolithization we studied the population genetic characteristics of the PN Aegean groups relative to central Anatolian groups. We first compared heterozygosity estimates among these populations. If the Aegeans were recent colonists from a single origin, due to a founder effect, one might expect lower heterozygosity in the Aegean than in central Anatolia. In contrast, Barcin and Revenia individuals had higher heterozygosity levels (mean 0.25 and 0.26, respectively) than those of Boncuklu and Tepecik (0.22 and 0.19, respectively) (**electronic supplementary material, table S7**).

Second, we calculated *D*-statistics focused on the Aegeans, which suggested higher admixture in this region than in central Anatolia:

(i) $D(\text{Yoruba}, \text{Natufian}; \text{—northAegean—}, \text{—centralAnatolia—}, \text{northAegean})$ revealed that pre-Neolithic population of Levant had stronger genetic affinity to the two north Aegean Neolithic populations (Barcin and Revenia, post-7,000 cal BCE), than to the two central Anatolian Neolithic groups (Boncuklu and

Tepecik, pre- and post-7,000 cal BCE) ($p < 0.05$; $Z \geq 3$) (**figure 2a, arrow “a”, electronic supplementary material, figure S4a, table S8**). Given the above-proposed gene flow event from Levant into Anatolia during the Neolithic this result might imply additional genetic interactions between Natufian-related populations and the ancestors of north Aegean populations that bypassed central Anatolia.

(ii) $D(\text{Yoruba}, \text{Caucasia/Iran}; \text{~~northAegean}~~, \text{centralAnatolia}, \text{northAegean})$ revealed that in 50% of the comparisons CHGs and Neolithic Iran individuals shared more alleles with the two north Aegean PN populations than with the two central Anatolians ($p < 0.05$; $Z \geq 2.8$) (**figure 2a, arrows “b” and “c”, electronic supplementary material, figure S4a, table S9**).

(iii) Likewise, WHG individuals showed higher affinity to the two north Aegean PN populations than PN central Anatolian group groups ($p < 0.05$; $Z \geq 3$) (**figure 2a, arrow “d”, electronic supplementary material, figure S4a, table S10**).

(iv) Natufians, WHGs and Iranian PPN individuals were consistently more similar to the Revenia individual than to those in Barcın (**electronic supplementary material, figure S4b, table S11**).

(v) Both the Boncuklu (PPN) and the Tepecik (PN) groups of central Anatolia had stronger affinity to the north Aegean PN populations, Barcın and Revenia, than to each other (**figure 2b, electronic supplementary material, table S12**). Likewise, all Anatolian groups (Boncuklu, Tepecik-Çiftlik and Barcın) were genetically closer to Revenia, than they were to each other (**figure 2b, electronic supplementary material, table S13**).

(vi) All European early farmer populations examined were genetically closer to Revenia than to each other (**electronic supplementary material, figure S4c, table S14**)

Observations (ii), (iii), and in particular (v) are intriguing. We asked whether these could be consistent with a number of demographic scenarios, assuming a phylogenetic topology that included Iran, WHG, and Aegean/Anatolian populations, estimated by [46]. We considered the following scenarios: (a) separate extreme bottlenecks in the ancestors of the two central Anatolian populations (possibly causing differentiation between the central Anatolian populations from each other, and from all other groups, (b) independent gene flow events from external sources (WHG and Iran) into the two central Anatolian groups (possibly causing differentiation between the two), (c) independent gene flow from WHG and Iran into the Aegean, (d) independent gene flow from WHG, Iran, and the two central Anatolian lineages into the Aegean. We performed serial coalescent simulations using realistic settings and compared the results with the observed *D*-statistics. We could only replicate the observed results under scenario (d) that describes rampant admixture in the Aegean (**electronic supplementary material, figure S5**).

4. Discussion

The analyses presented here highlight two points regarding the process of Neolithization. First, the observation that the two central Anatolian populations cluster together to the exclusion Neolithic populations of south Levant or of Iran restates the conclusion that farming in central Anatolia in the PPN was established by local groups instead of immigrants, which is consistent with the described cultural continuity between central Anatolian Epipaleolithic and Aceramic communities [9, 65]. This reiterates the earlier conclusion [33] that the early Neolithization in the primary zone was largely a process of cultural interaction instead of gene flow.

The second point relates to whether Aegean Neolithization (post-7,000 cal BCE) involved similar acculturation processes, or was driven by migration similar to Neolithization in mainland Europe - a long-standing debate in archaeology [16, 20, 22, 27, 28]. Here we discuss the two scenarios based on the genetic analysis.

Model 1: Migration from Anatolia to the Aegean. A recent study reported that by the 7th millenium BCE the eastward border of the WHG gene pool extended to the Iron Gates (on the border between Romania and Serbia) [34]. Plausibly, during the early Holocene, the WHG population could also have been present along the Aegean coastline, such that the border between central Anatolian and WHG gene pools ran along west Anatolia. If so, the Aegean Neolithization must have involved replacement of a local, WHG-related Mesolithic population by incoming easterners.

If migration occurred, where did it originate? Because Revenia and Barcın cluster with PPN and PN central Anatolian Neolithic groups to the exclusion of the south Levant (**figure 1c, electronic supplementary material, figure S3c**), the latter is unlikely to be the source, leaving central Anatolia or south Anatolia (north Levant) as potential origins.

Notably, the north Aegeans (Revenia and Barcın) show higher diversity than the central Anatolians. We had earlier shown that the highest quality Barcın genome carries a smaller proportion of short runs of homozygosity than the highest quality Boncuklu genome [31], which also supports the notion that the ancestral effective population size of the Aegeans was larger than those of central Anatolians. Moreover, we find that the north Aegeans share more alleles with eastern, western and southern gene pools, as estimated using the *D*-statistic (**figure 2**). Although the *D*-statistic can be sensitive to technical biases, our result is unlikely to be a technical artifact because, (a) the north Aegean data were derived from two independent studies [32, 38], (b) the Barcın data was produced using two different techniques, whole genome shotgun sequencing and SNP capture, and (c) both Barcın and Revenia display the same population genetic patterns, suggesting that the admixture signals in the Aegean individuals are reproducible. In addition, although unknown population structure can complicate interpretation of the *D*-statistic [64], we note that the admixture estimates are consistent with the estimated higher genetic diversity in the Aegean.

If the Revenia and Barcın individuals studied here were descendants of Anatolian Neolithic immigrants, they must have been recent settlers, as all samples analyzed here date to early stages of the Aegean Neolithic (Revenia: 6,438-6,264 and Barcın: 6,500-6,200 cal BCE). Furthermore, if the migration was *directly* of central Anatolian origin (represented by Boncuklu and Tepecik-Çiftlik), the putative migrants must have admixed with populations carrying alleles of distinct gene pools (Levant, Caucasus/Iran, and WHG) within a few centuries, in order to explain our observations above (**figure 2a**).

Alternatively, the migration event could have originated from the Anatolian south coast or north Levant [27] (currently no genome data is available from these groups). This region could have hosted a hypothetical central Anatolian-related population exposed to admixture from CHG-, Iran-, and Levant-related gene pools in earlier millennia. A south Anatolian population could have been in contact with different central Anatolian populations from the Konya Plain (Boncuklu) and Cappadocia (Tepecik-Çiftlik), explaining the affinity of both Boncuklu and Tepecik-Çiftlik to Barcın. A seafaring population could also be in genetic contact with putative WHG-related populations of the Aegean. This hypothetical population could have initiated the Cyprus Neolithic in the 11th millennium BCE and later Aegean Neolithic communities in the 7th millennium BCE [27].

One surprising observation here is the apparent absence of WHG-like ancestry in Late Neolithic/Chalcolithic Aegean genomes: ADMIXTURE analysis results from two individuals from northwest Anatolia (Kumtepe, ~5000 BCE) [74] and four individuals from south Greece (Franchti Cave and Diros, ~4000 BCE) [34], all lack noticeable WHG-like ancestry components [31,32,34]. This contrasts with WHG admixture emerging in European farmer populations in the Middle and Late Neolithic [37, 39], and perhaps earlier in the Balkans [34], indicating the persistence of Mesolithic populations in Europe after Neolithic migrations. Therefore, if the Mesolithic populations of the Aegean coast had indeed been WHG-related, they must have been fully replaced by the eastern migrant farmers.

Model 2: Adoption of Neolithic elements by local foragers. Alternatively, the Aegean coast Mesolithic populations may have been part of the Anatolian-related gene pool that occupied the Aegean seaboard during the early Holocene. Under this scenario, the north Aegean PN populations would be at least partial descendants of local hunter-gatherers who adopted Neolithic lifestyle post-7,000 cal BCE, triggered by contacts with central Anatolian and Levantine populations. The following events would be conceivable: (a) During the Last Glacial Maximum (LGM), the Aegean evolved into a refuge hosting a significant human population, which is in line with climatic modeling [66-69]; estimates of human population density during the Marine Isotope Stage 2 in west (but not central) Anatolia reach one of their highest levels in Europe [70, 71]. The existence of an Aegean human population going back to the LGM is also consistent with mitochondrial haplogroup-based analyses [72], and that Anatolian-like mitochondrial haplogroups are found also in Mesolithic Balkan and Aegean populations [32, 34]. (b) Following the LGM, Aegean emigrants dispersed into central Anatolia and established populations that eventually gave rise to the local Epi-Paleolithic and later Neolithic communities, in line with the earliest direct evidence for human presence in central Anatolia c.14,000 cal BCE [61]. This hypothetical out-of-the-Aegean event coincides with the post-LGM Near East-related migration signatures in European Mesolithic genomes [73]. (c) Between the LGM and post-7,000 cal BCE Neolithization, WHG, Natufian, and Caucasus/Iran-related groups admixed with north Aegeans, differentiating the latter from their central Anatolian relatives and leading to our observations in **figure 2a**. (d) Post-7,000 cal BCE, there occurred additional, albeit limited central Anatolian gene flow back into the Aegean, giving rise to our observation in **figure 2b**.

The archaeological evidence. Both the migration and acculturation models for Aegean Neolithisation enjoy support from material culture investigations, but the overall evidence points to a complex process where Aegean societies were culturally influenced by diverse sources, including the central Anatolian Neolithic, the Levant Neolithic, and possibly local Mesolithic traditions. In contrast to the relative homogeneity of European Neolithic cultures, such as the LBK and Cardial, the Aegean Neolithic is noted for its diversity [65]. Variation in Neolithic Package elements and primary zone traditions is notable across Aegean sites,

among regions (e.g. east and west of Marmara), even between closely neighbouring villages [16,17,20,65,75-80]. This diversity includes, for example, obsidian, with Greek Aegean (Melos) [81,82] or mainland Anatolian (Cappadocian) [83] sources being preferred in some settlements, and yet other settlements showing no evidence of obsidian use [65]. Cultural trait diversity involves architecture, tool types, ceramics, and symbolic elements (such as figurines and intramural burial), which may show partial similarities to either central Anatolia or to the Levant, or may be unique [16,65,76]. For instance, intramural burial, a common feature among primary zone sites, is also widespread in east Marmara early Neolithic villages (including Barcın), but totally absent in settlements only two hundred kilometers west [16]. Mesolithic-like lithic industries and the prominence of seafood in some settlements further imply the continuing presence of Aegean Mesolithic traditions into the Neolithic [16,20,27,84,85]. Indeed, lively seafaring activity was prevalent in the Mediterranean and the Aegean already by the 11th millennium BCE [86,87], as evidence from Cyprus, Crete, Franchti, Cyclops Cave, Ouirakos and other Aegean island and coastal mainland Mesolithic sites demonstrate [20,27,84,87-95].

Instead of a single-sourced colonisation process, the Aegean Neolithization may thus have flourished upon already existing coastal and interior interaction networks connecting Aegean foragers with Levantine and central Anatolian PPN populations, and involved multiple cultural interaction events from its early steps onward [16,20,65,75]. This wide diversity of cultural sources and the potential role of local populations in Neolithic development may set apart Aegean Neolithization from that in mainland Europe. While Mesolithic Aegean genetic data is awaited to fully resolve this issue, researchers should be aware of the possibility that the initial emergence of the Neolithic elements in the Aegean, at least in the north Aegean, involved cultural and demographic dynamics different than those in European Neolithization.

Acknowledgements:

We thank T.Günther, A.Munters, S.C.Acan, A.Birand, Y.S.Erdal, A.Fairbairn, A.Omrak, M.Jakobsson, M.Krzewinska, F.Özer, R.Özbal, L.Excoffier, and J.Stora for technical support in analysis, suggestions and/or comments on the manuscript, and two anonymous reviewers for helpful suggestions. E.A. were supported by the Institutional Development Program of the University of Ostrava and by EU structural funding Operational Programme Research and Development for Innovation, project No. CZ.1.05/2.1.00/19.0388

Funding:

TÜBİTAK (grant no. 114Z927 to M.S.), TÜBA GEBİP (M.S.), METU BAP (to M.S.), Knut and Alice Wallenberg Foundation: "1000 ancient genomes" (to A.G.).

Contributions:

M.S., I.T., G.M.K., A.G., Ç.A. designed and supervised the study; G.M.K., D.K., A.P.S. analysed data with contributions from H.M.D., R.Y., E.A., P.F.; Ç.A., D.B., A.M.B., C.C.B. provided archaeological and ecological interpretations; M.S., I.T., G.M.K., Ç.A. wrote the manuscript with input from all authors.

References:

1. Kelly RL. 1995 *The foraging spectrum: diversity in hunter-gatherer lifeways*. Smithsonian Institution Press.
2. Zeder MA. 2009 Evolutionary biology and the emergence of agriculture: The value of co-opted models of evolution in the study of culture change. In *Macroevolution in Human prehistory*, pp. 157–210. New York, NY: Springer New York. (doi:10.1007/978-1-4419-0682-3_7)
3. Aiello LC. 2011 The Origins of agriculture: New data, New ideas. *Curr. Anthropol.* **52**, S161–S162. (doi:10.1086/660154)
4. Özdoğan M. 1997 The Beginning of Neolithic economies in Southeastern Europe: an Anatolian perspective. *J. Eur. Archaeol.* **5**, 1–33.

- 429 5. Bar-Yosef O. 2001 From sedentary foragers to village hierarchies: the emergence of social
430 institutions. *Proc. Br. Acad.* **101**, 1–38.
- 431 6. Kuijt I, Finlayson B. 2009 Evidence for food storage and predomestication granaries 11,000 years
432 ago in the Jordan Valley. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 10966–70.
433 (doi:10.1073/pnas.0812764106)
- 434 7. Zeder MA. 2012 The Broad spectrum revolution at 40: Resource diversity, intensification, and an
435 alternative to optimal foraging explanations. *J. Anthropol. Archaeol.* **31**, 241–264.
436 (doi:10.1016/j.jaa.2012.03.003)
- 437 8. Weiss E, Zohary D. 2011 The Neolithic Southwest Asian founder crops. *Curr. Anthropol.* **52**,
438 S237–S254. (doi:10.1086/658367)
- 439 9. Baird D. 2012 The Late Epipaleolithic, Neolithic, and Chalcolithic of the Anatolian plateau,
440 13,000-4000 BC. In *A companion to the archaeology of the ancient Near East*, pp. 431–465.
441 Oxford, UK: Wiley-Blackwell. (doi:10.1002/9781444360790.ch23)
- 442 10. Kuijt I, Goring-Morris N. 2002 Foraging, farming, and social complexity in the pre-Pottery
443 neolithic of the southern Levant: A Review and Synthesis. *J. World Prehistory* **16**, 361–440.
444 (doi:10.1023/A:1022973114090)
- 445 11. Akkermans PMMG, Schwartz GM. 2003 *The archaeology of Syria: from complex hunter-*
446 *gatherers to early urban societies (c. 16,000-300 BC)*. Cambridge University Press.
- 447 12. Flannery K V. 2002 The Origins of the village revisited: From nuclear to extended households.
448 *Am. Antiq.* **67**, 417–433. (doi:10.2307/1593820)
- 449 13. Asouti E, Fuller DQ. 2013 A Contextual approach to the emergence of agriculture in southwest
450 Asia. *Curr. Anthropol.* **54**, 299–345. (doi:10.1086/670679)

- 451 14. Atakuman Ç. 2014 Architectural discourse and social transformation during the early Neolithic of
452 southeast Anatolia. *J. World Prehistory* **27**, 1–42. (doi:10.1007/s10963-014-9070-4)
- 453 15. Özdoğan M. 2011 Archaeological evidence on the westward expansion of farming communities
454 from eastern Anatolia to the Aegean and the Balkans. *Curr. Anthropol.* **52**, S415–S430.
455 (doi:10.1086/658895)
- 456 16. Özdoğan M. 2014 A new look at the introduction of the Neolithic way of life in Southeastern
457 Europe. Changing paradigms of the expansion of the Neolithic way of life. *Doc. Praehist.* **41**, 33–
458 49.
- 459 17. Perles C. 2005 From the Near East to Greece: Let's reverse the Focus – Cultural elements that did
460 not transfer. In *how did farming reach Europe?: Anatolian-European relations from the second*
461 *half of the 7th through the first half of the 6th millennium cal BC* (ed C Lichter), pp. 275–290.
462 İstanbul: Byzas 2.
- 463 18. Zvelebil M. 1989 On the transition to farming in Europe, or what was spreading with the
464 Neolithic: a reply to Ammerman (1989). *Antiquity* **63**, 379–383.
- 465 19. Zvelebil M. 2009 Mesolithic societies and the transition to farming: problems of time, scale and
466 organisation. In *Hunters in Transition: Mesolithic Societies of Temperate Eurasia and Their*
467 *Transition to Farming*, Cambridge University Press.
- 468 20. Reingruber A. 2011 Early Neolithic settlement patterns and exchange networks in the Aegean.
469 *Doc. Praehist.* **38**, 291. (doi:10.4312/dp.38.23)
- 470 21. Brami M, Zanotti A. 2015 Modelling the initial expansion of the Neolithic out of Anatolia. *Doc.*
471 *Praehist.* **42**, 103. (doi:10.4312/dp.42.6)
- 472 22. Çilingiroğlu Ç, Çakırlar C. 2013 Towards configuring the neolithisation of Aegean Turkey. *Doc.*
473 *Praehist.* **40**, 21–29. (doi: 10.4312/dp.40.3)

- 474 23. Ammerman AJ, Cavalli-Sforza LL (Luigi L. 1984 *The neolithic transition and the genetics of*
475 *populations in Europe*. Princeton: Princeton University Press.
- 476 24. Renfrew C. 1990 *Archaeology and language : the puzzle of Indo-European origins*. Cambridge
477 University Press.
- 478 25. Özdoğan M. 2007 Von Zentralanatolien nach Europa. In *Vor 12000 Jahren in Anatolien. Die*
479 *ältesten Monumente der Menschheit*, p. 160. Stuttgart: Badisches Landesmuseum Karlsruhe.
- 480 26. Özdoğan M. 2008 An alternative approach in tracing changes in demographic composition.
481 *Neolithic demographic transition and its consequences*, 139–178.
- 482 27. Horejs B, Milić B, Ostmann F, Thanheiser U, Weninger B, Galik A. 2015 The Aegean in the early
483 7th millennium BC: Maritime networks and colonization. *J. World Prehistory* **28**, 289–330.
484 (doi:10.1007/s10963-015-9090-8)
- 485 28. Zvelebil M. 2001 The agricultural transition and the origins of Neolithic society in Europe. *Doc.*
486 *Praehist.* **28**, 1-26. (doi: 10.4312/dp.28.1)
- 487 29. Zvelebil M. 1989 On the transition to farming in Europe, or what was spreading with the
488 Neolithic: a reply to Ammerman. *Antiquity* **63**, 379–383.
- 489 30. Alasdair W. R. Whittle. 1996 *Europe in the Neolithic: The creation of New Worlds*. Cambridge
490 University Press.
- 491 31. Kılınç GM *et al.* 2016 The demographic development of the first farmers in Anatolia. *Curr. Biol.*
492 **26**, 2659–2666. (doi:10.1016/j.cub.2016.07.057)
- 493 32. Hofmanová Z *et al.* 2016 Early farmers from across Europe directly descended from Neolithic
494 Aegeans. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 6886–91. (doi:10.1073/pnas.1523951113)
- 495 33. Lazaridis I *et al.* 2016 Genomic insights into the origin of farming in the ancient Near East. *Nature*
496 **536**, 419–424. (doi:10.1038/nature19310)

- 497 34. Mathieson I *et al.* 2017 The Genomic History of Southeastern Europe. *BioRxiv*
498 (doi:10.1101/135616)
- 499 35. Jones ER *et al.* 2017 The Neolithic transition in the Baltic was not driven by admixture with early
500 European Farmers. *Curr. Biol.* **27**, 1–7. (doi:10.1016/J.CUB.2016.12.060)
- 501 36. Gamba C *et al.* 2014 Genome flux and stasis in a five millennium transect of European prehistory.
502 *Nat. Commun.* **5**, 5257. (doi:10.1038/ncomms6257)
- 503 37. Günther T *et al.* 2015 Ancient genomes link early farmers from Atapuerca in Spain to modern-day
504 Basques. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 11917–22. (doi:10.1073/pnas.1509851112)
- 505 38. Mathieson I *et al.* 2015 Genome-wide patterns of selection in 230 ancient Eurasians. *Nature* **528**,
506 499–503. (doi:10.1038/nature16152)
- 507 39. Haak W *et al.* 2015 Massive migration from the steppe was a source for Indo-European languages
508 in Europe. *Nature* **522**, 207–211. (doi: 10.1038/nature14317)
- 509 40. Skoglund P, Malmström H, Raghavan M, Storå J, Hall P, Willerslev E, Gilbert MTP, Götherström
510 A, Jakobsson M. 2012 Origins and genetic legacy of Neolithic farmers and hunter-gatherers in
511 Europe. *Science*. **336**.
- 512 41. Skoglund P *et al.* 2014 Genomic diversity and admixture differs for Stone-Age Scandinavian
513 foragers and farmers. *Science* **344**, 747–750. (doi:10.1126/science.1253448)
- 514 42. Bollongino R *et al.* 2013 2000 years of parallel societies in Stone Age central Europe. *Science*
515 **342**, 479–481. (doi: 10.1126/science.1245049)
- 516 43. Lazaridis I *et al.* 2014 Ancient human genomes suggest three ancestral populations for present-day
517 Europeans. *Nature* **513**, 409–413. (doi:10.1038/nature13673)
- 518 44. Olalde I *et al.* 2014 Derived immune and ancestral pigmentation alleles in a 7,000-year-old
519 Mesolithic European. *Nature* **507**, 225–228. (doi:10.1038/nature12960)

- 520 45. Gallego-Llorente M *et al.* 2016 The genetics of an early Neolithic pastoralist from the Zagros,
521 Iran. *Sci. Rep.* **6**, 31326. (doi:10.1038/srep31326)
- 522 46. Broushaki F *et al.* 2016 Early Neolithic genomes from the eastern Fertile Crescent. *Science* (80-.).
523 **353**, 499–503. (doi:10.1126/science.aaf7943)
- 524 47. Jones ER *et al.* 2015 Upper Palaeolithic genomes reveal deep roots of modern Eurasians. *Nat.*
525 *Commun.* **6**, 8912. (doi:10.1038/ncomms9912)
- 526 48. Li H, Durbin R. 2009 Fast and accurate short read alignment with Burrows-Wheeler transform.
527 *Bioinformatics* **25**, 1754–1760. (doi:10.1093/bioinformatics/btp324)
- 528 49. Kircher M. 2012 Analysis of high-throughput ancient DNA sequencing data. In *Methods in*
529 *molecular biology* (Clifton, N.J.), pp. 197–228. (doi:10.1007/978-1-61779-516-9_23)
- 530 50. Patterson N *et al.* 2012 Ancient admixture in human history. *Genetics* **192**, 1065–1093.
531 (doi:10.1534/genetics.112.145037)
- 532 51. McVean GA *et al.* 2012 An integrated map of genetic variation from 1,092 human genomes.
533 *Nature* **491**, 56–65. (doi:10.1038/nature11632)
- 534 52. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R.
535 2009 The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079.
536 (doi:10.1093/bioinformatics/btp352)
- 537 53. Danecek P *et al.* 2011 The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158.
538 (doi:10.1093/bioinformatics/btr330)
- 539 54. Purcell S *et al.* 2007 PLINK: a tool set for whole-genome association and population-based
540 linkage analyses. *Am. J. Hum. Genet.* **81**, 559–75. (doi:10.1086/519795)
- 541 55. Patterson N *et al.* 2006 Population structure and eigenanalysis. *PLoS Genet.* **2**, e190.
542 (doi:10.1371/journal.pgen.0020190)

- 543 56. Benjamini Y and Yekutieli D. 2001 The control of the false discovery rate in multiple testing
544 under dependency. *The Annals of Statistics* **29**, 1165-1188.
- 545 57. Quinlan AR, Hall IM. 2010 BEDTools: a flexible suite of utilities for comparing genomic features.
546 *Bioinformatics* **26**, 841–842. (doi:10.1093/bioinformatics/btq033)
- 547 58. Korneliussen TS, Albrechtsen A, Nielsen R. 2014 ANGSD: Analysis of next generation
548 sequencing data. *BMC Bioinformatics* **15**, 356. (doi:10.1186/s12859-014-0356-4)
- 549 59. Reich D. *et al.* 2012 Reconstructing Native American population history. *Nature* **488**, 370–374.
550 (doi:10.1038/nature11258)
- 551 60. Excoffier, L. and M. Foll. 2011 fastsimcoal: a continuous-time coalescent simulator of genomic
552 diversity under arbitrarily complex evolutionary scenarios. *Bioinformatics* **27**: 1332- 1334.
- 553 61. Baird D *et al.* 2013 Juniper smoke, skulls and wolves’ tails. The Epipalaeolithic of the Anatolian
554 plateau in its South-west Asian context; insights from Pınarbaşı. *Levant* **45**, 175–209.
- 555 62. Bıçakçı E, Godon M, Çakan YG. 2012 Tepecik-Çiftlik. In *Neolithic In Turkey: New excavations*
556 *and new research 3–Central Turkey* (eds M Özdoğan, N Başgelen, P Kuniholm), pp. 89–134.
557 İstanbul: Archaeology and Art Publications.
- 558 63. Alpaslan Roodenberg S, Gerritsen FA, Özbal R. 2013 Neolithic burials from Barcın Höyük: The
559 2007-2012 Excavation Seasons. *Anatolica* **39**, 93–111.
- 560 64. Eriksson A and Manica A. 2012 Effect of ancient population structure on the degree of
561 polymorphism shared between modern human populations and ancient hominins. *Proc. Natl.*
562 *Acad. Sci. U. S. A.* **109**, 13956-13960. (doi: 10.1073/pnas.1200567109)
- 563 65. Düring B. 2011 *The prehistory of Asia Minor, from complex hunter-gatherers to early urban*
564 *societies*. Cambridge: Cambridge University Press.

- 565 66. Şenkul Ç, Doğan U. 2013 Vegetation and climate of Anatolia and adjacent regions during the Last
566 Glacial period. *Quat. Int.* **302**, 110–122. (doi:10.1016/j.quaint.2012.04.006)
- 567 67. Sarıkaya MA, Çiner A. 2015 Late Pleistocene glaciations and paleoclimate of Turkey. *Bull. Miner.*
568 *Res. Explor.* **0**, 107–127. (doi:10.19111/bmre.35245)
- 569 68. Shumilovskikh LS, Tarasov P, Arz HW, Fleitmann D, Marret F, Nowaczyk N, Plessen B, Schlütz
570 F, Behling H. 2012 Vegetation and environmental dynamics in the southern Black Sea region
571 since 18kyr BP derived from the marine core 22-GC3. *Palaeogeogr. Palaeoclimatol. Palaeoecol.*
572 **337**, 177–193. (doi:10.1016/j.palaeo.2012.04.015)
- 573 69. Valsecchi V, Sanchez Goñi MF, Londeix L. 2012 Vegetation dynamics in the Northeastern
574 Mediterranean region during the past 23 000 yr: insights from a new pollen record from the Sea of
575 Marmara. *Clim. Past* **8**, 1941–1956. (doi:10.5194/cp-8-1941-2012)
- 576 70. Timmermann A, Friedrich T. 2016 Late Pleistocene climate drivers of early human migration.
577 *Nature* **538**, 92–95. (doi:10.1038/nature19365)
- 578 71. Giampoudakis K, Marske KA, Borregaard MK, Ugan A, Singarayer JS, Valdes PJ, Rahbek C,
579 Nogués-Bravo D. 2017 Niche dynamics of Palaeolithic modern humans during the settlement of
580 the Palaearctic. *Glob. Ecol. Biogeogr.* **26**, 359–370. (doi:10.1111/geb.12543)
- 581 72. Pereira JB *et al.* 2017 Reconciling evidence from ancient and contemporary genomes: a major
582 source for the European Neolithic within Mediterranean Europe. *Proc. R. Soc. London B Biol. Sci.*
583 **284**.
- 584 73. Fu Q *et al.* 2016 The genetic history of Ice Age Europe. *Nature* **534**, 200–5.
585 (doi:10.1038/nature17993)
- 586 74. Omrak A *et al.* 2016 Genomic evidence establishes Anatolia as the source of the European
587 Neolithic gene pool. *Curr. Biol.* **26**, 270–5. (doi:10.1016/j.cub.2015.12.019)

- 588 75. Yakar J. 2016 The nature and extent of Neolithic Anatolia's contribution to the emergence of
589 farming communities in the Balkans: an overview. In *south east Europe and Anatolia in prehistory*
590 *Band 263: Essays in honor of Vassil Nikolov on His 65th Anniversary* (eds K Bacvarov, R Gleser),
591 pp. 25–68. Bonn: Verlag Dr. Rudolf Habelt GmbH.
- 592 76. Düring BS. 2013 Breaking the Bond: Investigating The Neolithic expansion in Asia Minor in the
593 seventh millennium BC. *J. World Prehistory* **26**, 75–100. (doi:10.1007/s10963-013-9065-6)
- 594 77. Thissen L. 2011 The Neolithic–Chalcolithic sequence in the SW Anatolian Lakes Region. *Doc.*
595 *Praehist.* **37**, 269. (doi:10.4312/dp.37.23)
- 596 78. Borić D. 2007 Mesolithic-Neolithic interactions in the Danube Gorges. In *Mesolithic/Neolithic*
597 *interactions in the Balkans and in the middle Danube Basin* (eds JK Kozłowski, M Nowak), pp.
598 31–45. Oxford.
- 599 79. Fuller DQ, Willcox G, Allaby RG. 2011 Cultivation and domestication had multiple origins:
600 arguments against the core area hypothesis for the origins of agriculture in the Near East. *World*
601 *Archaeol.* **43**, 628–652. (doi:10.1080/00438243.2011.624747)
- 602 80. Bailey DW. 2000 *Balkan prehistory: exclusion, incorporation and identity*. Routledge.
- 603 81. Perlès C, Takaoğlu T, Gratuze B. 2011 Melian obsidian in NW Turkey: Evidence for early
604 Neolithic trade. *J. F. Archaeol.* **36**, 42–49. (doi:10.1179/009346910X12707321242313)
- 605 82. Milić M. 2014 PXRF characterisation of obsidian from central Anatolia, the Aegean and central
606 Europe. *J. Archaeol. Sci.* **41**, 285–296. (doi:10.1016/j.jas.2013.08.002)
- 607 83. Çilingiroğlu A, Çilingiroğlu Ç. 2007 Ulucak. In *Anadolu'da uygarlığın doğuşu Avrupaya yayılımı*
608 *Türkiye'de Neolitik dönem, yeni kazılar, yeni bulgular* (eds M Özdoğan, N Başgelen), pp. 361–
609 372. İstanbul: Archaeology and Art Publications.

- 610 84. Efstratiou N, Biagi P, Karkanas E, Starnini E. 2013 A Late Palaeolithic site at Ouriakos (Limnos,
611 Greece) in the north-eastern Aegean. *Antiquity*. Project Gallery. See
612 <http://antiquity.ac.uk/projgall/efstratiou335/>.
- 613 85. Yalçinkaya I *et al.* 1995 Les occupations tardiglaciaires du site d'Öküzini (Sudouest de la Turquie)
614 Résultats préliminaires. *L'Anthropologie* **99**, 562–583.
- 615 86. Broodbank C. 2006 The origins and early development of Mediterranean maritime activity. *J.*
616 *Mediterr. Archaeol.* **19**. (doi:10.1558//jmea.2006.v19i2.199)
- 617 87. Vigne J-D, Zazzo A, Saliège J-F, Poplin F, Guilaine J, Simmons A. 2009 Pre-Neolithic wild boar
618 management and introduction to Cyprus more than 11,400 years ago. *Proc. Natl. Acad. Sci. U. S.*
619 *A.* **106**, 16135–8. (doi:10.1073/pnas.0905015106)
- 620 88. Peltenburg EJ, Wasse A. 2004 *Neolithic revolution: new perspectives on southwest Asia in light of*
621 *recent discoveries on Cyprus*. Oxford: Oxbow Books. See
622 <http://www.oxbowbooks.com/dbbc/neolithic-revolution.html>.
- 623 89. Vigne J-D *et al.* 2012 First wave of cultivators spread to Cyprus at least 10,600 y ago. *Proc. Natl.*
624 *Acad. Sci. U. S. A.* **109**, 8445–9. (doi:10.1073/pnas.1201693109)
- 625 90. Bar-Yosef O. 2015 Southeast Turkish Neolithic: A view from the southern Levant. In *The*
626 *Neolithic in Turkey 6* (eds M Özdoğan, N Başgelen, P Kuniholm), pp. 293–320. İstanbul.
- 627 91. Takaoğlu T, Korkut T, Erdoğan B, Işin G. 2014 Archaeological evidence for 9th and 8th millennia
628 BC at Girmeler cave near Tlos in SW Turkey. *Doc. Praehist.* **41**, 111–118. (doi:10.4312/dp.41.6)
- 629 92. Simmons AH. 2007 *The neolithic revolution in the near east: Transforming the human landscape*.
630 University of Arizona Press.
- 631 93. Sampson A. 2005 New evidence from the early pottery production stages in the Aegean Basin
632 from the 9th to the 7th millennium cal BCE. In *How did farming reach Europe?: Anatolian-*

European relations from the second half of the 7th through the first half of the 6th millennium cal BC (ed C Lichter), pp. 131–141. İstanbul: Byzas 2.

94. Sampson A. 2008 *The cave of the cyclops: Mesolithic and Neolithic networks in the northern Aegean, Greece. Volume I: Intra-Site analysis, local industries, and regional site distribution*. Institute for Aegean Prehistory Academic Press.

95. Sampson A, editor. 2011 *The cave of the cyclops: Mesolithic and Neolithic networks in the northern Aegean, Greece, Volume II: Bone tool industries, dietary resources and the paleoenvironment, and archeometrical studies*. INSTAP Academic Press.

Figure 1. Genetic differentiation among ancient west Eurasians and predicted admixture events. Panels (a) and (c) show results of the same multidimensional scaling (MDS) analysis, summarising f_3 -statistics (shared genetic drift) between ancient population pairs (electronic supplementary material, table S2). The goodness of fit was estimated as 0.17 and 0.17 for both dimensions. Admixture events among gene pools inferred using D -statistics are represented as arrows on each MDS plot. The circles where the arrow tips touch indicate which population is involved in the inferred admixture. Tepecik-Çiftlik is labeled as Tepecik. **(a)** Admixture in Boncuklu (central Anatolian PPN). For clarity, the other Anatolian/Aegean populations are not plotted. Arrow “a”: gene flow between Boncuklu and pre-Neolithic populations of mainland Europe (relative to other gene pools). Arrow “b”: gene flow between Boncuklu and Levant populations (relative to other gene pools). Arrow “c”: gene flow between Boncuklu and Caucasia/Iran populations (relative to other gene pools). **(b)** Results of D -statistics in the form of $D(\text{Yoruba}, p1; p2, \text{Boncuklu})$. Multiple testing correction was performed using the Benjamini-Yekutieli method [56]. **(c)** Arrow “d”: estimated gene flow between Boncuklu and Levantine populations. This is based on testing the topology $D(\text{Outgroup}, \text{Boncuklu}; \text{Levant}_{\text{preNeolithic}}, \text{Levant}_{\text{Neolithic}})$, showing that the Boncuklu population showed higher genetic affinity to Levantine Neolithics (sample ages: c.8,300-6,750 cal BCE) than to Levantine pre-Neolithics (sample ages: c.11,840-9,760 cal BCE), although the result was only

marginally significant ($p > 0.05$, $Z > 2.5$) (figure S2c, table S5). Arrow “e”: gene flow from the Iran PPN population into Anatolian/Aegean PN populations. Arrow “f”: gene flow from the Levant PPN population into Anatolian/Aegean PN populations (electronic supplementary material, table S5 and table S6).

Figure 2. Summary of D -statistics describing population relationships within the Anatolian/Aegean gene pool and between Anatolians/Aegeans and neighboring groups. The Yoruba genome was used as outgroup in D -statistics. The Tepecik-Çiftlik is labeled as Tepecik. All D -statistics results are reported in electronic supplementary material, table S8-S13. **(a)** D -statistics results summarised as arrows on the MDS plot (same as figure 1). Each triple population compared in D -tests are framed in the same colour. If a test population has greater genetic affinity to the second population compared to a third one, an arrow with same color as the frames is drawn from the test population to the second population (the arrows' direction or lengths are not representative of gene flow magnitudes). Arrow “a” and navy frames summarise $D(\text{Yoruba}, \text{Natufian}; \text{northAegean}, \text{centralAnatolian}, \text{northAegean})$, where Natufians had stronger genetic affinity to north Aegean PN than to central Anatolian PPN or PN groups (electronic supplementary material, table S8). Arrows “b” and “c” and green frames summarise $D(\text{Yoruba}, \text{CHG\&Iran_Neolithic}; \text{northAegean}, \text{centralAnatolian})$. In 6/8 comparisons CHGs and Iran PPN populations had stronger genetic affinity to the north Aegean PN than to central Anatolian PPN and PN (electronic supplementary material, table S9). Arrow “d” and purple frames summarise $D(\text{Yoruba}, \text{WHG}; \text{northAegean}, \text{centralAnatolian})$. In all comparisons WHGs had stronger genetic affinity to the north Aegean PN than to central Anatolian PPN and PN, with the exception of $D(\text{Yoruba}, \text{WHG}; \text{Boncuklu}, \text{Barcın})$ being non-significant (electronic supplementary material, table S10). **(b)** Results of D -tests calculated as $D(\text{Outgroup}, \text{RightPopulation}; \text{BottomPopulation}, \text{LeftPopulation})$, where right, bottom and left refer to the positions of the populations on the matrix. For instance the top row shows that Boncuklu has significantly higher affinity to Barcın than to Tepecik-Çiftlik. The D -statistic magnitude is represented by color, Z score by size, and significance by being filled or not. Multiple testing correction

was performed using the Benjamini-Yekutieli method [56] (electronic supplementary material, table S12 and S13)

Figure 1

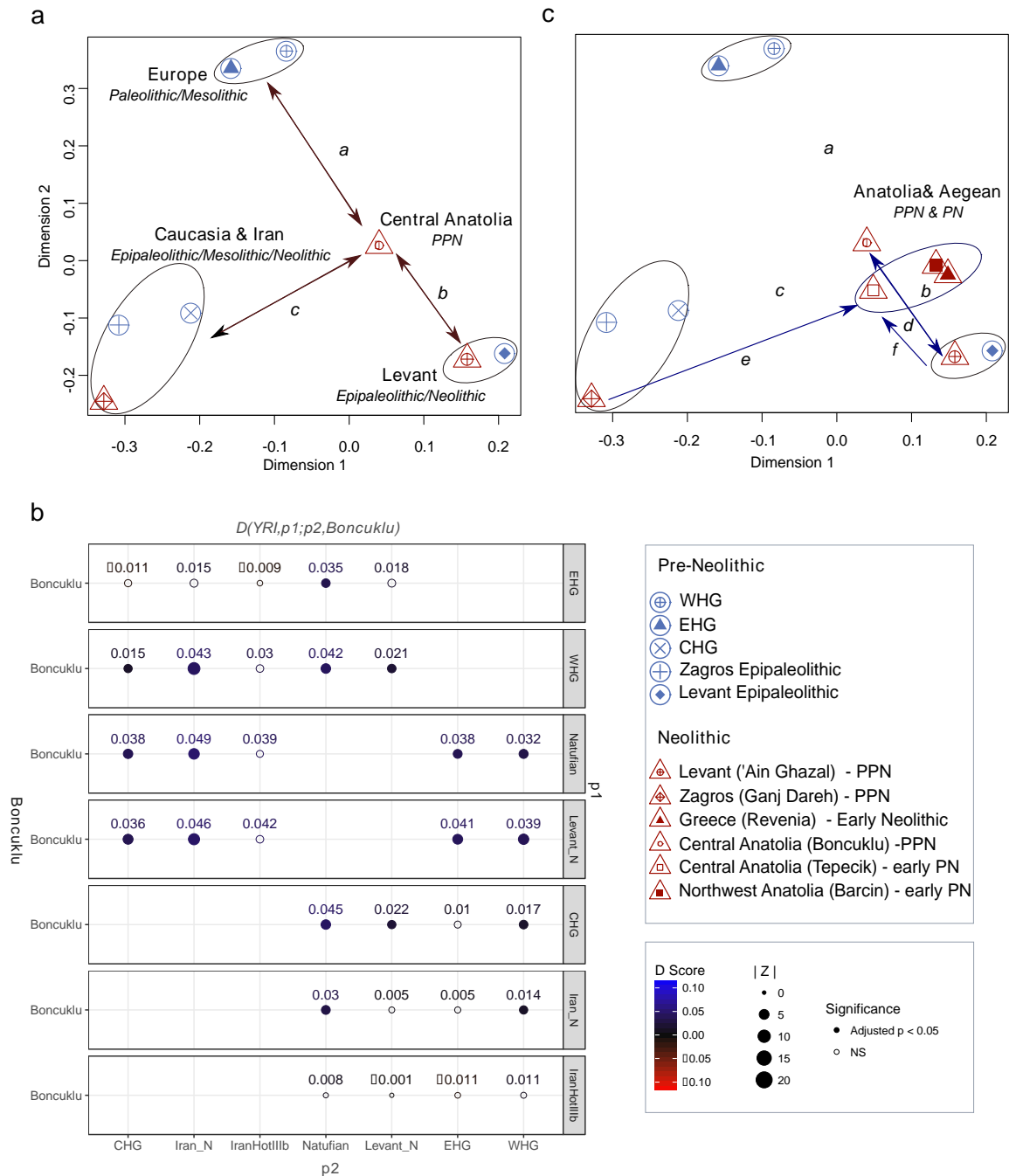


Figure 1. Genetic differentiation among ancient west Eurasians and predicted admixture events. Panels a and c show the results of the same multidimensional scaling (MDS) analysis, summarising f_3 -statistics (shared genetic drift) between ancient population pairs (electronic supplementary material, table S2). The goodness of fit was estimated as 0.17 and 0.17 for two dimensions. Admixture events among gene pools inferred using D -statistics are represented as arrows on each MDS plot. The circles where the arrow tips touch indicate which population is involved in the inferred admixture. Tepecik-Çiftlik is labeled as Tepecik. **(a)** Admixture in Boncuklu (central Anatolian PPN). For clarity, the other Anatolian/Aegean populations are not plotted. Arrow “a”: gene flow between Boncuklu and pre-Neolithic populations of mainland Europe (relative to other gene pools). Arrow “b”: gene flow between Boncuklu and Levant populations (relative to other gene pools). Arrow “c”: gene flow between Boncuklu and Caucasia/Iran populations (relative to other gene pools). **(b)** Results of D -statistics in the form of $D(\text{Yoruba}, p1; p2, \text{Boncuklu})$. Multiple testing correction was performed including all D -statistics results discussed in the paper using Benjamini-Yekutieli method [56]. **(c)** Arrow “d”: estimated gene flow between Boncuklu and Levantine populations. This is based on testing the topology $D(\text{Outgroup}, \text{Boncuklu}; \text{Levant_preNeolithic}, \text{Levant_Neolithic})$, showing that the Boncuklu population showed higher genetic affinity to Levantine Neolithics (sample ages: c.8,300-6,750 cal BCE) than to Levantine pre-Neolithics (sample ages: c.11,840-9,760 cal BCE), although the result was only marginally significant ($p > 0.05$, $Z > 2.5$) (figure S2c, table S5). Arrow “e”: gene flow from the Iran PPN population into Anatolian/Aegean PN populations. Arrow “f”: gene flow from the Levant PPN population into Anatolian/Aegean PN populations (electronic supplementary material, table S5 and table S6).

Figure 2

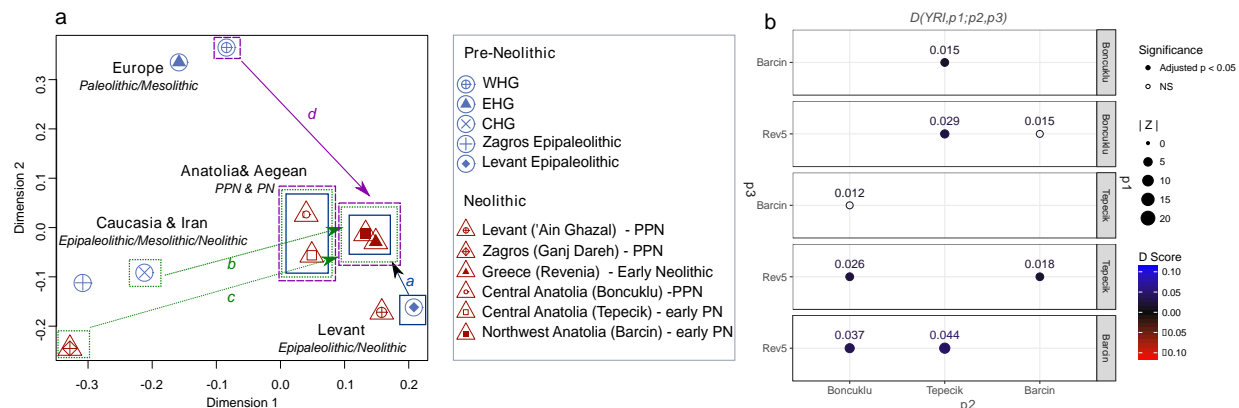


Figure 2. Summary of D -statistics describing population relationships within the Anatolian/Aegean gene pool and between Anatolians/Aegeans and neighboring groups. The Yoruba genome was used as outgroup in D statistics. The Tepecik-Çiftlik is labeled as Tepecik. All D statistics results are reported in electronic supplementary material, table S8-S13. (a) D -statistics results summarised as arrows on the MDS plot (same as figure 1). Each triple population compared in D tests are framed in the same colour. If a test population has greater genetic affinity to the second population compared to a third one, an arrow with same color as the frames is drawn from the test population to the second population (the arrows' direction or lengths are not representative of gene flow magnitudes). Arrow "a" and navy frames summarise $D(\text{Yoruba}, \text{Natufian}; \text{northAegean}, \text{centralAnatolian})$, where Natufians had stronger genetic affinity to north Aegean PN than to central Anatolian PPN or PN groups (electronic supplementary material, table S8). Arrows "b" and "c" and green frames summarise $D(\text{Yoruba}, \text{CHG\&Iran_Neolithic}; \text{northAegean}, \text{centralAnatolian})$. In 6/8 comparisons CHGs and Iran PPN populations had stronger genetic affinity to the north Aegean PN than to central Anatolian PPN and PN (electronic supplementary material, table S9). Arrow "d" and purple frames summarise $D(\text{Yoruba}, \text{WHG}; \text{northAegean}, \text{centralAnatolian})$. In all comparisons WHGs had stronger genetic affinity to the north Aegean PN than to central Anatolian PPN and PN, with the exception of $D(\text{Yoruba}, \text{WHG}; \text{Boncuklu}, \text{Barcin})$ being non-significant (electronic supplementary material, table S10). (b) Results of D tests calculated as $D(\text{Outgroup}, \text{RightPopulation}; \text{BottomPopulation}, \text{LeftPopulation})$. Multiple testing correction was performed including all D -statistics results discussed in the paper using Benjamini-Yekutieli method (electronic supplementary material, table S12 and S13).

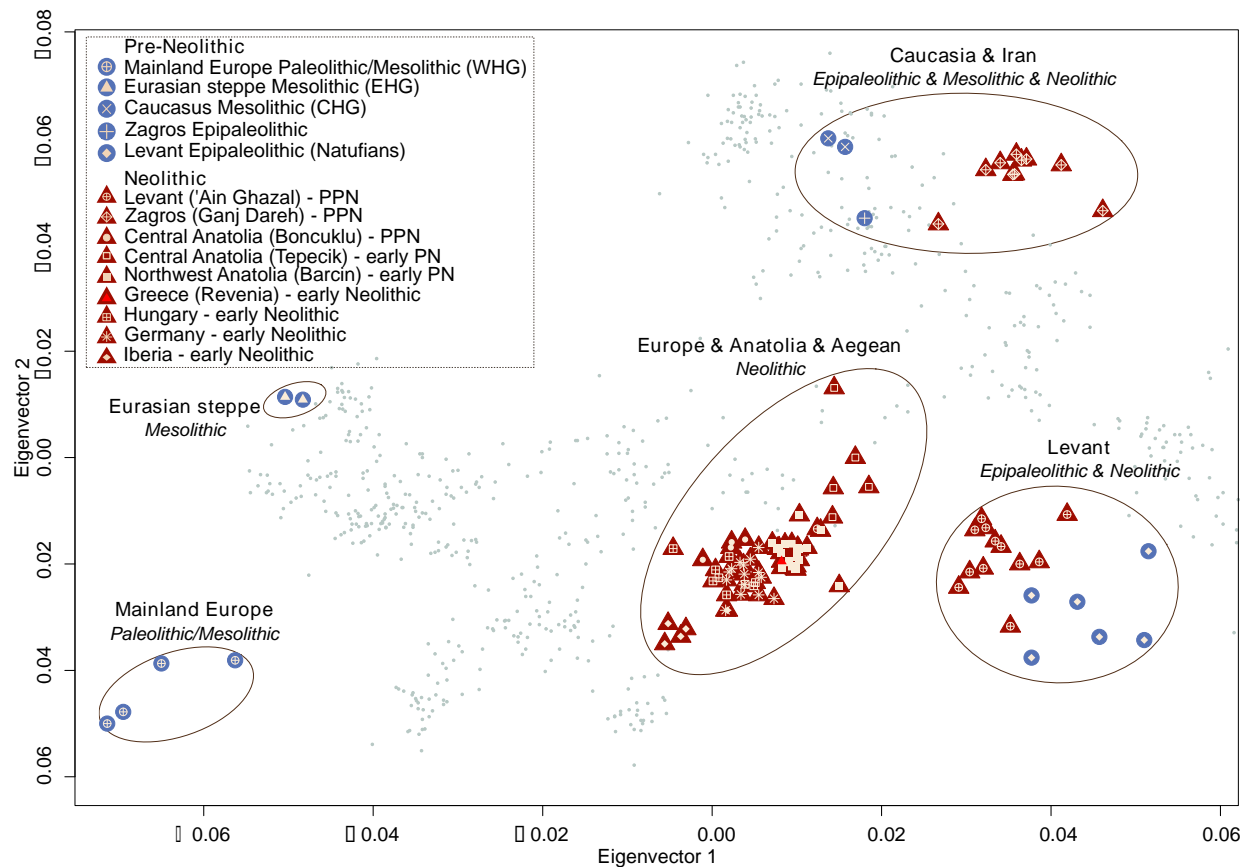


Figure S1. Principal component analysis (PCA) with modern and ancient genomes. The eigenvectors were calculated using 50 modern west Eurasian populations, onto which genome data from ancient individuals were projected. The gray circles highlight the four ancient gene pools of west Eurasia. Modern-day individuals are shown as gray points. In the Near East, Pre-Neolithic (Epipaleolithic/Mesolithic) and Neolithic individuals genetically cluster by geography rather than by cultural context. For instance, Neolithic individuals of Anatolia cluster to the exclusion of individuals from the Levant or Iran). In Europe, genetic clustering reflects cultural context but not geography: European early Neolithic individuals are genetically distinct from European pre-Neolithic individuals but tightly cluster with Anatolians. PPN: Pre-Pottery/Aceramic Neolithic, PN: Pottery Neolithic, Tepecik: Tepecik-Çiftlik (electronic supplementary material, table S1 lists the number of SNPs per ancient individual).

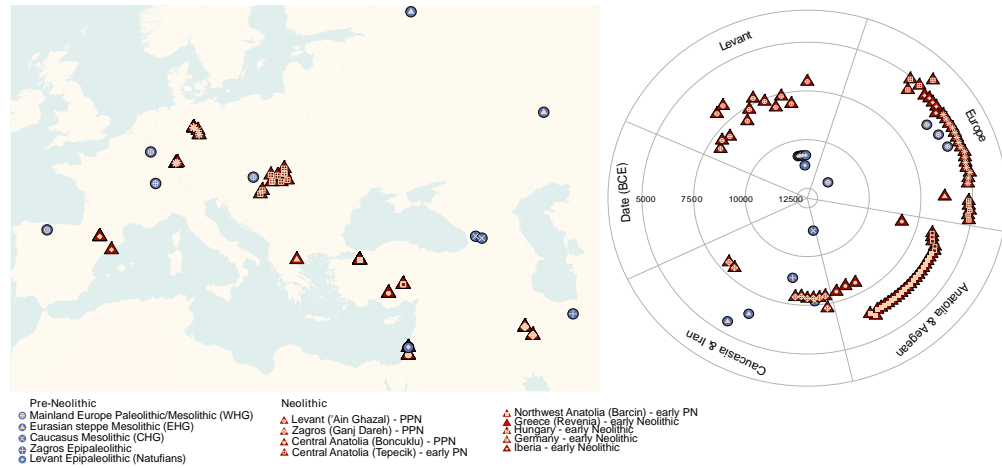


Figure S2. Summary of the data analyzed in this study. (a) Map of west Eurasia showing the geographical locations and (b) timeline showing the time period (years BCE) of ancient individuals investigated in the study. Blue circles: individuals from pre-Neolithic context; red triangles: individuals from Neolithic contexts. For further information about the samples see electronic supplementary material, table S1.

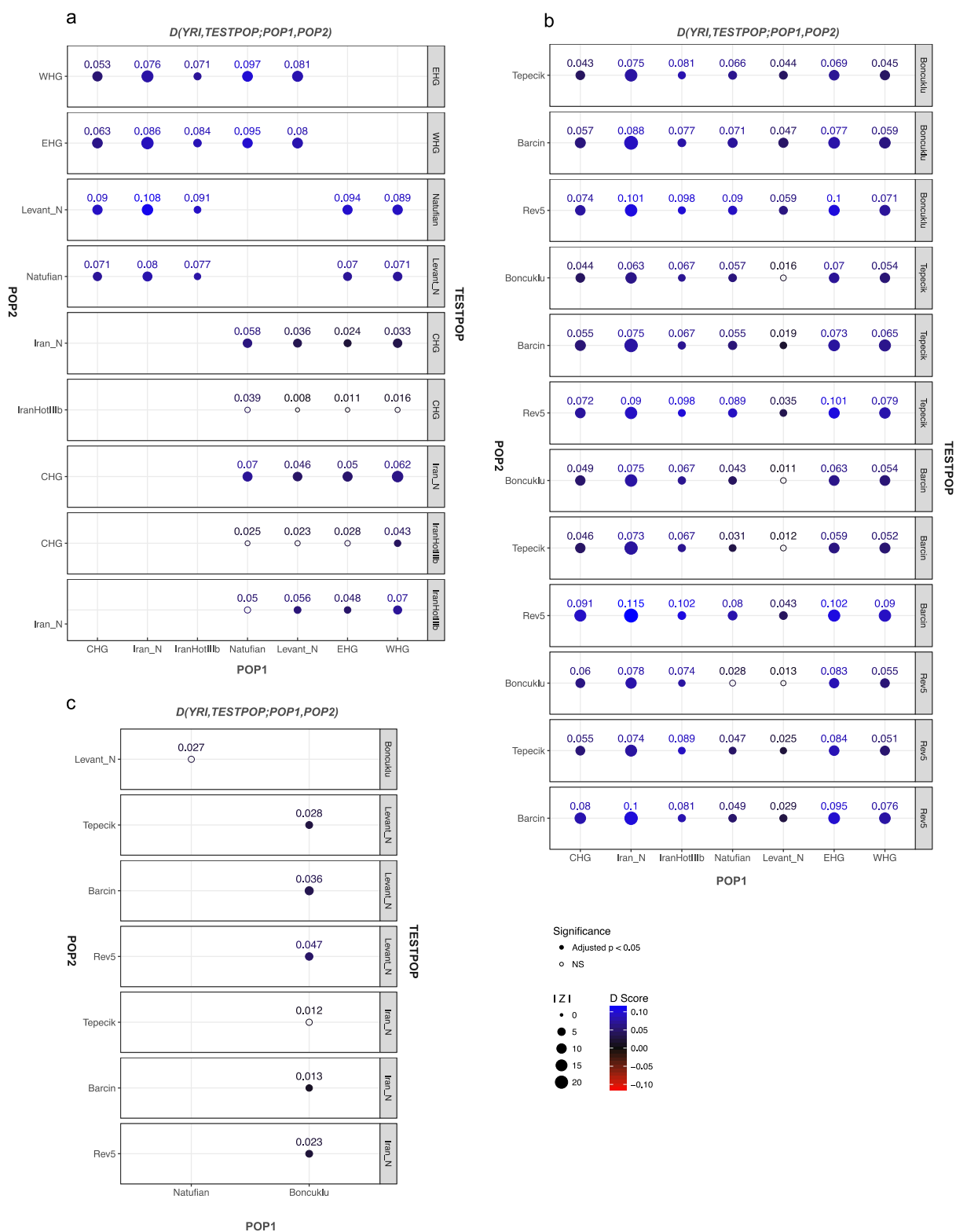


Figure S3. Plots showing the results of D -statistics with topology $D(\text{Yoruba}, \text{right population}, \text{bottom population}, \text{left population})$. Multiple testing correction was performed

including all D -statistics results discussed in the paper using Benjamini-Yekutieli method [56]. Results of D -statistics in the form of (a) $D(\text{Yoruba}, p1; p2, p3)$. “ p ” refers to the Caucasasia/Iran, the Levant, or European pre-Neolithic gene pools (Table S3). (b) $D(\text{Yoruba}, p1; p2, p3)$. “ $p1$ ” and “ $p2$ ” are Anatolian/Aegean populations and “ $p3$ ” is either Caucasasia/Iran, Levant, or European pre-Neolithic (Table S5). (c) $D(\text{Yoruba}, p1; p2, p3)$. “ $p1$ ” is Iran or Levant Neolithic, “ $p2$ ” is Boncuklu and $p3$ is one of the PNs. (Table S6)

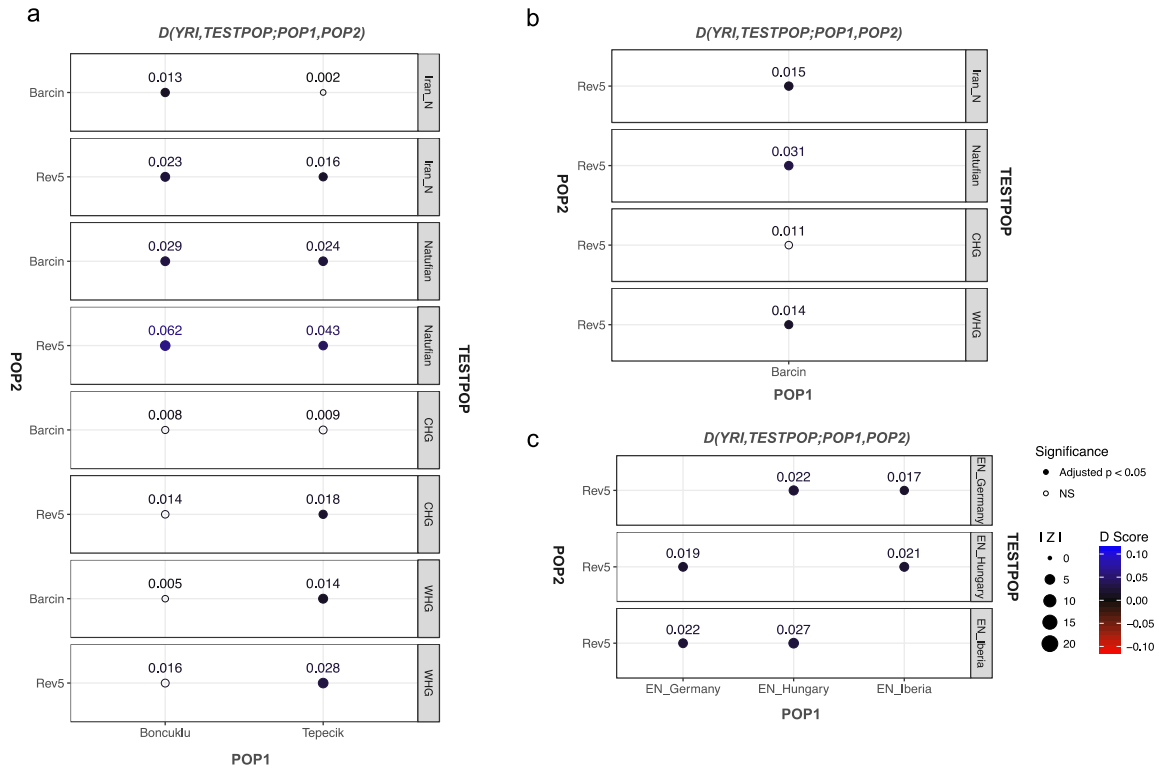


Figure S4. Plots showing the results of D -statistics with topology $D(\text{Yoruba}, \text{right population}, \text{bottom population}, \text{left population})$. Multiple testing correction was performed including all D -statistics results discussed in the paper using Benjamini-Yekutieli method [56]. Results of D -statistics in the form of (a) $D(\text{Yoruba}, \text{Natufian}; \text{northAegean}, \text{centralAnatolia})$ (Table S8-S10). (b) $D(\text{Yoruba}, \text{Iran Neolithic}\&\text{Natufian}\&\text{CHG}\&\text{WHG}; \text{Barcin}, \text{Revenia})$ (Table S11). (c) $D(\text{Yoruba}, \text{European Neolithic}; \text{European Neolithic}, \text{Revenia})$ (Table S14)